



Southeast Alaska Network

SEAN Kittlitz's Murrelet Database User Guide

Abstract

Observations and related data from Kittlitz's murrelet monitoring surveys in Glacier Bay are maintained by SEAN in a consistent database. This paper summarizes the scope of these data. Methods for accessing data are explained. References are given for finding definitions of the collected attributes.

Table of Contents

1	The Origin of These Data	3
2	Availability of Detailed Data	3
3	Known Data Limitations	3
4	Mechanics of Accessing Data	3
5	Spatial Coverage.....	4
6	Temporal Coverage	4
7	Authority	4
8	Analysis Facilities	5
9	Further Information	5
10	Description of Data Fields.....	5

Revision history

Revision Date	Base Version	Author	Nature of Change
7/17/2013	4/18/2013	Bill Johnson	Corrected typos
4/18/2013	-new-	Bill Johnson	Initial version

1 The Origin of These Data

SEAN has been performing annual surveys in order to estimate Kittlitz's murrelet abundance and density in Glacier Bay since 2010. Details regarding the nature of these surveys, the statistical framework, and the formal data specifications are available in the monitoring protocol. This and other documentation is available at SEAN's web site:

http://science.nature.nps.gov/im/units/sean/KM_Main.aspx.

Three sets are available for download from the monolithic database. Deliverable KM_A, transect definitions, is a table listing names, locations, sampling frequency, and statistical parameters of the 88 permanent transects identified for the survey. KM_E contains the complete observation data set. This includes trackpoints recorded every two seconds during field work, showing murrelet species, group size, behavior, and angle to centerline when encountered. KM_K is a table of abundance estimates and associated statistical parameters, calculated using distance sampling models.

Additional data products, while not housed under a DBMS, are available in a variety of forms, as defined in the protocol's table 4.1.

2 Availability of Detailed Data

Data collection was begun in calendar year 2009 under a draft protocol, eventually named KM-2012.1. Considerable changes were made to collecting methods after the first year was examined and significant improvements were implemented. The consistent SEAN observation data series is available covering 2010 and beyond.

SEAN's data are in the public domain.

3 Known Data Limitations

There are currently no known limitations on the collected data. Exceptional circumstances are occasionally encountered when attempting to follow a very strict protocol under field conditions. Such exceptions are documented in field observer notes, available in PDF form as deliverable KM_B.

4 Mechanics of Accessing Data

Three authoritative database-housed products are available on the "Original Source Data" tab of the SEAN web site at http://science.nature.nps.gov/im/units/sean/KM_Main.aspx. The purpose of this web page is to provide access to the basic data. Results are available as both screen grids and file downloads. There is a button on the web page to generate each. Grids are populated with results and may be copied and pasted into other applications using standard Windows facilities. The grid is generally limited to 1,000 rows. This limit protects users from overloading their

computers; there are potentially many rows available, each with many columns. In actual operation, it is usually convenient to adjust the filtering parameters until the grid contents confirm that desired data are being delivered. Then data sets of any size may be saved on one's workstation in comma separated value files by pressing the "Save Results to a CSV File" button.

Filtering may be done by specifying values in the appropriate fields on the page. Filter parameters restrict the result set. If filter parameters are omitted, then all values are included for that attribute. Note that the KM_A transect deliverable has no filter. The other two allow entry of various screening values.

A particular year may be selected, or all years. If a set of, say, two or three specific years are desired, then a separate download must be made for each. Alternatively, all years may be selected which then may be narrowed down by specifying a range of start and end dates that cross a span of multiple years.

A sampling date range may be specified for KM_E. If earliest date is omitted, the first date ever sampled is used. If latest date is omitted, then today's date is used. Be aware that specifying a date range lying outside the chosen year will result in finding zero rows.

A check box is available for KM_E where one may specify whether or not to include comment fields in the extract. Because comments on each row can be lengthy, it is sometimes more convenient to retrieve data without them.

5 Spatial Coverage

The monitoring program uses a set of specially selected line transects to represent Glacier Bay proper. This deliberately omits the outer coast of Glacier Bay National Park and Preserve. Also, a few closed wilderness areas are not accessible to survey boats, so their populations are not represented. The unrepresented areas are defined in the protocol document.

The transect scheme is considered fixed; it is not expected to ever change.

6 Temporal Coverage

SEAN always surveys in July. Its estimates do not purport to be representational of abundance at other times of year.

7 Authority

The data browser pulls rows directly from SEAN's vital sign database. These data reflect the authoritative source. Final analyses should always be based on data pulled from the browser. Other sources, such as spreadsheet files, paper tables typed back into computers, or email communications cannot be relied on to be authoritative.

SEAN is in the process of arranging for external repositories to house these data. Such repositories include NPS's IRMA Data Store. It should be noted that the maintenance of data copied to these repositories is out of the control of SEAN. If any data discrepancy is seen among the sources, the SEAN website should be considered the authority.

8 Analysis Facilities

The data browsing feature gives everyone the ability to extract data of interest onto their local computers. From there they are free to use whatever tools they choose to perform whatever analyses they are interested in. The data browser does not provide embedded analytical tools.

9 Further Information

Additional information, as well as current contact persons, may be found in the Kittlitz's murrelet section of <http://science.nature.nps.gov/im/units/sean/>.

10 Description of Data Fields

Detailed explanations of every field are provided in the protocol document itself, available on SEAN's main Kittlitz's murrelet web page. Note that each row in the database is stamped with the protocol it was collected and validated under, and different protocol versions may treat specific fields somewhat differently.

Over time, expect new protocol designators to appear in order to reflect updates to the monitoring program methods. Be sure to select the matching protocol document when investigating the fine details of data fields.